

## **CLAIMS**

### **What is claimed is:**

1. A method for speech enabling an application comprising the steps of:  
specifying a speech input with a speech-enabled markup;  
defining within said speech-enabled markup at least one operation of an application that is to be executed upon a detection of said specified speech input;  
after said defining step, instantiating said application;  
detecting said specified speech input; and  
executing said application operation responsive to said detecting step.
2. The method of claim 1, wherein said application is a multimodal Web browser.
3. The method of claim 1, further comprising the steps of:  
providing a speech-enabled markup interpreter within an operating system upon which said application executes, wherein said speech-enabled markup interpreter is used to detect said speech input and responsively initiate said application operation.
4. The method of claim 3, further comprising the steps of:  
rendering a Web page within said application, wherein said Web page includes speech-enabled markup for at least one element of said Web page, and wherein said speech-enabled markup interpreter speech-enables said Web page element.
5. The method of claim 1, further comprising the steps of:  
associating said speech-enabled markup with a graphical user interface element of said application;  
determining that said graphical user interface element receives focus; and  
responsive to said determination, activating said speech-enabled markup so that said application starts monitoring audible input for said specified speech input.
6. The method of claim 5, further comprising the steps of:  
determining that said graphical user interface element loses focus; and

responsive to said loss of focus, deactivating said speech-enabled markup so that said application no longer monitors audible input for said specified speech input.

7. The method of claim 1, wherein said application is written in a Markup language.

8. The method of claim 1, wherein said speech-enabled markup is written in a Voice Extensible Markup Language.

9. The method of claim 8, wherein said application is written in an Extensible Hypertext Markup Language.

10. A speech-enabled application comprising:

a graphical user interface element configured to initiate at least one application operation responsive to a predefined graphical user interface event;

a speech-enabled markup associated with said graphical user interface element that specifies said application operation is to be performed responsive to a speech input; and

a markup interpreter configured to interpret said speech-enabled markup and initiate said application operation responsive to said speech input.

11. The application of claim 10, wherein said markup interpreter is embedded within an operating system of a client computer in which said application is disposed.

12. The application of claim 10, wherein said speech-enabled application is a Web browser.

13. The application of claim 12, wherein said markup interpreter is configured to interpret speech-enabled markup contained within Web pages rendered by said Web browser.

14. The application of claim 10, wherein said application is written in a markup language.

15. A machine-readable storage having stored thereon, a computer program having a plurality of code sections, said code sections executable by a machine for causing the machine to perform the steps of:

- specifying a speech input with a speech-enabled markup;
- defining within said speech-enabled markup at least one operation of an application that is to be executed upon a detection of said specified speech input;
- after said defining step, instantiating said application;
- detecting said specified speech input; and
- executing said application operation responsive to said detecting step.

16. The machine-readable storage of claim 15, wherein said application is a multimodal Web browser.

17. The machine-readable storage of claim 15, further comprising the steps of:

- providing a speech-enabled markup interpreter within an operating system upon which said application executes, wherein said speech-enabled markup interpreter is used to detect said speech input and responsively initiate said application operation.

18. The machine-readable storage of claim 17, further comprising the steps of:

- rendering a Web page within said application, wherein said Web page includes speech-enabled markup for at least one element of said Web page, and wherein said speech-enabled markup interpreter speech-enables said Web page element.

19. The machine-readable storage of claim 15, further comprising the steps of:

- associating said speech-enabled markup with a graphical user interface element of said application;
- determining that said graphical user interface element receives focus; and

responsive to said determination, activating said speech-enabled markup so that said application starts monitoring audible input for said specified speech input.

20. The machine-readable storage of claim 19, further comprising the steps of:  
determining that said graphical user interface element loses focus; and  
responsive to said loss of focus, deactivating said speech-enabled markup so that said application no longer monitors audible input for said specified speech input.

21. The machine-readable storage of claim 15, wherein said application is written in a Markup language.

22. The machine-readable storage of claim 15, wherein said speech-enabled markup is written in a Voice Extensible Markup Language.

23. The machine-readable storage of claim 22, wherein said application is written in an Extensible Hypertext Markup Language.

24. A system for speech enabling an application comprising:  
means for specifying a speech input with a speech-enabled markup;  
means for defining within said speech-enabled markup at least one operation of an application that is to be executed upon a detection of said specified speech input;  
means for instantiating said application after said defining step;  
means for detecting said specified speech input; and  
means for executing said application operation responsive to said detecting step.